HbA1c

In this assessment, you need to create a full report for the descriptive analysis of the dataset provided. After proper data cleaning and formatting, you need to decide what statistics to present and how to present them. The report MUST have, at least, one figure and one table. There is no limit for the number of tables or figures. Your text should be similar of a results section of a proper scientific paper. No statistical testing or modelling is required or desired.

The report MUST be created using RMarkdown and the file submitted needs to be the .rmd file ONLY (html final format). As with assessment 2, no effort will be made to adapt parts of the script. It must run perfectly and the result of it will be assessed.

As part of the assessment, you need to prepare a 5-10min presentation to be presented in class (in person or online) in week 13 (01/06/2022).

This simulated dataset shows a sample 1234 participants to investigate the effect of a treatment on the glycated hemoglobin level (HbA1c). The data dictionary is presented below.

Variable	Description	Codes
id	Participant ID	
Age	Age	numeric
Sex	Sex	0 – Female 1 – Male
Group	Group allocation	0 – Control 1 – Treatment
Height	Participant height	numeric
Weight	Participant weight	numeric
HbA1c_Pre	Participant HbA1c measure before the study	numeric
HbA1c_Post	Participant HbA1c measure after the study	numeric

Variables to be created
Age in categories: break the age into two categories (up to 55, 56 or more)
BMI: calculate the body mass index (BMI) using weight/(height/100)^2
BMI in categories: break the BMI into four categories (Low – up to 18.5, Normal – >18.5 and <25, Overweight – 25-29.9, Obese – 30 or more)